DOCUMENT RESUME

ED 064 350 TM 001 517

AUTHOR Stake, Robert E.
TITLE An Approach to the Evaluation of Instructional Programs.
PUB DATE Apr 72
NOTE 8p.; Paper presented at the annual meeting of the American Educational Research Association (Chicago, Illinois, April 1972)

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS *Evaluation Criteria; *Evaluation Techniques; *Instructional Programs; *Program Content; *Program Evaluation

ABSTRACT

The evaluation of an educational program by portrayal of the program rather than the focus of the program is discussed. It is suggested that the program evaluator limit his evaluation aims to what he can do and to what the client needs most. It is believed that the first duty of the evaluator should be to offer the client a comprehensive portrayal of the program. (DB)

FILMED FROM BEST AVAILABLE COPY

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

ED 064350

AN APPROACH TO THE EVALUATION OF INSTRUCTIONAL PROGRAMS*

Robert E. Stake
University of Illinois at Urbana-Champaign

The evaluator of an instructional program is faced with a dilemma. His design and final report can emphasize what he can measure most effectively given his modest resources--or his design and final report can eflect the nature of the program, with fidelity to the many important perceptions and expectations of it. Both cannot prevail. What the evaluator has to say cannot be both a sharp analysis of high-priority achievement and a broad and accurate reflection of the program's complex transactions.

I am saying something more than: "You can't feature both product and process in the evaluation study." I am saying, "Any focus on the analysis of product or process distorts the picture as to what the instruction is. Which is more important: to tell of some very special things about the program or to provide the most veridical portrayal of the program?

TM 001 512

* Paper delivered at AERA Annual Meeting April 4, 1972, in Chicago.

I am today going to advocate the latter, the emphasis on portrayal, at the expense of focus. I do not mean this to be an emphasis on descriptive data rather than judgmental; both are needed in veridical portrayal evaluations and in sharp-focus evaluations. I do not mean this to be an emphasis on summative rather than formative evaluation; in either case the same dilemma appears. I acknowledge that any study that emphasizes a particular issue, or a particular decision, or a particular goal, at the great dimunition of all others, might be a most appropriate research or evaluation effort--but it should not be passed off as an evaluation <u>of the program</u>.

<u>Our Limited Talents</u>.

We recognize that we are not equally able to measure all outcomes. Some community expectations and student aspirations are deeply hidden. Some costs are more difficult to tease out than others. Working with a limited evaluation budget we are inclined to confine our attention to that which we can measure best.

Many of us have a great confidence that we can measure any trait, describe any event, and operationalize any construct though we are quick to allow that we cannot do those jobs with a paper-and-pencil-test,

with a one-day site visit, or with an annual budget of $2,000. In fact, many of the promises we make in evaluation proposals could not be fulfilled if we had a full-time evaluation team consisting of the ten most recent past Presidents of AERA, at least half of whom are considered experts in evaluation. But the point at the moment is not that evaluation is a tough job but that some of our tasks in any program evaluation are tougher than others. And there is a reasonable tendency on our part to feature in the evaluation proposal and periodic reports those tasks we do best.

What jobs do we do best? We are inclined to say that the jobs we do best are those our fellow researchers admire or least criticize: our item analyses, sociograms, task analyses, random samplings, covariance analyses, attitude scales, mail surveys, and so on. And so we suggest to our clients that those things would be useful to them. We sometimes imply that the evaluation would not be authentic if it did not use some of them. We fail to realize that many of these procedures were brought into our technology as microscopes to examine the minute detail of education, not as procedures for portraying the "whole cloth" of an instructional program.

The Multiplicity of Program Components

The whole cloth of an educational program is a grand accumulation of intents, transactions, and outcomes. The teachers intend to deliver on many promises and to take advantage of many targets of opportunity. Students and parents have their expectations and apprehensions. Community leaders, social critics, and educationists have "viewed with alarm" and "pointed with pride." Each child brings his own complex of convictions, misunderstandings, and propensities and takes away some of those and still others. Each classroom is a community, with rules and stresses and competition and compassion. Yesterday's subgroups are not tomorrow's. Things are learned, unlearned, relearned much as shoelaces are knotted, untied, broken, and retied. An educational program has countless objectives, many of them dormant until a crisis arises. The priorities vary over time from person to person. No statement of program objectives ever devised has come close to representing the real-world intents of the people involved in an educational program.

I should not imply that one cannot get reasonable consensus on high-priority objectives for a program. The unspoken objectives—safety in the classroom; sharing of work responsibilities; developing a

sense of humor, a respect for rules, and a tolerance of ambiguity; and so on, and so on—the unspoken objectives are left to take care of themselves, at least until a crisis arises. And then these objectives may pre empt all others. One can get simple consensus; and as long as no one takes the consensus too seriously, children can get much more than the primitive education that the consensus statement describes.

Consensus is one of the great simplifiers. Theory is another. Statistical processes are simplifiers. Test scores are simple representations of the complex. These simplifiers help us by reducing the phenomena to something within our power of comprehension. But they mislead us by saying that education is much less than it really is. We work day by day with the simplifications—the statements of objectives, the central tendencies, the criterion tests—and we become transfixed by them, losing our awareness of the fundamental activities of teaching and learning. We do it to ourselves and we do it to our audiences. Evaluators should be helping people keep in touch with the reality of instruction, but our scrapbooks are full of enlargements of enlargements.

Should the evaluator focus on the more prominent features or attempt to embrace the program as a whole? There are many different

demands on the evaluator. Let me quote from a notice to potentential contractors from the U.S. Office of Education for evaluation of the Experimental Schools Program:

> "Contractors will be required to implement a comprehensive evaluation of the project including the following areas: (1) measurement of the success of the overall project in meeting needs of students and of the impact of the project; (2) measurement of individual project elements; (3) systematic documentation of the project; (4) description of the social and political forces which influence and shape the development of the project; (5) formative evaluation for the project and its internal evauluation efforts; and (6) cost analysis."

All these evaluation goals are important. But the annual bill for a single school's comprehensive evaluation--as defined--would run over a million dollars. That's probably one reason why no evaluation team has ever succeeded in doing all those things. In the more ordinary and modest situation we should try to be comprehensive too, but we should limit our evaluation aims to what we can do and to what the client needs most. What many clients need is a credible, thorough representation of what the program is, including information about who likes what about it. Clients need confirmation of what they know, reminders of things they are overlooking, and something in the way of a report to show other people.

It is difficult for the client to perceive the scope and movement of the program. The program director's perspective is partially

obscured; the outsider's is evanescent. They need to see more, to share more in the experience. If the program glows, the evaluation should reflect some of it. If the program wobbles, the tremor should pass through the evaluation report. The first duty of the evaluator should be to offer the client a comprehensive portrayal of the program.

The client may want something else. O.K. He may want more than portrayal. He may want something other than portrayal. O.K. He may want a concentrated examination of the pursuit of a few objectives. He may want a study of the causes of success or failure, or a study of transportability, or a study of the efficiency of the program. If he has the resources, he should get what he wants. But he should not be encouraged to pursue those costly and elusive phantoms if what he needs is a substantive portrayal of his instructional program.

It's a tough choice: focus or portrayal. The evaluator has to help figure out which will be more useful. Many of us are biased in favor of focus.

Think for a moment what a book review has become in the Sunday Times: an opportunity for the reviewer to get something off his

chest, a chance to pamper a pet idea, with at most a tenuous connection to the book reviewed. Are program evaluation studies connected to programs by more than a tenuous shoestring? Are they little more than the exploiting of an instructional researcher's hunch or a psychometrician's fascination? We owe the people more than that.